Human Molecular Genetics, 2004, Vol. 13, No. 24 **3189–3202**
doi:10.1093/hmg/ddh334
Advance Access published on October 20, 2004

Branch site haplotypes that control alternative splicing

Jana Královičová[1], Sophie Houngninou-Molango[2], Angela Krämer[2] and Igor Vořechovský[1,*]

[1]University of Southampton School of Medicine, Division of Human Genetics, Southampton SO16 6YD, UK and [2]University of Geneva, Department of Cell Biology, CH-1211 Geneva 4, Switzerland

Received August 25, 2004; Revised October 6, 2004; Accepted October 12, 2004 GenBank accession numbers: AY553273–AY553292

We show that the allele-dependent expression of transcripts encoding soluble HLA-DQβ chains is determined by branchpoint sequence (BPS) haplotypes in *DQB1* intron 3. BPS RNAs associated with low inclusion of the transmembrane exon in mature transcripts showed impaired binding to splicing factor 1 (SF1), indicating that alternative splicing of *DQB1* is controlled by differential BPS recognition early during spliceosome assembly. We also demonstrate that naturally occurring human BPS point mutations that alter splicing and lead to recognizable phenotypes cluster in BP and in position −2 relative to BP, implicating impaired SF1–BPS interactions in disease-associated BPS substitutions. Coding DNA variants produced smaller fluctuations of exon inclusion levels than random exonic substitutions, consistent with a selection against coding mutations that alter their own exonization. Finally, proximal splicing in this multi-allelic reporter system was promoted by at least seven SR proteins and repressed by hnRNPs F, H and I, supporting an extensive antagonism of factors balancing the splice site selection. These results provide the molecular basis for the haplotype-specific expression of soluble DQβ, improve prediction of intronic point mutations and indicate how extraordinary, selection-driven DNA variability in HLA affects pre-mRNA splicing.

INTRODUCTION

Alternative pre-mRNA splicing (AS), or a generation of multiple transcripts from a single pre-mRNA, is a fundamental mechanism of regulating gene expression and generating proteomic diversity (1). The estimated frequency of human genes that undergo AS has risen dramatically from ∼5% to well over 50% in the last decade (2), but the proportion of the AS events that have functional consequences is currently unknown (3).

The recognition of splice sites involves cross-talk between multiple interactions that contribute to complexes that commit the pre-mRNA to splicing. The splicing reaction occurs in the spliceosome, which consists of the pre-mRNA substrate, four small nuclear ribonucleoprotein particles (snRNPs U1, U2, U4/U6 and U5) and a large number of proteins (4). In higher eukaryotes, the requirement for accurate splicing is accompanied by weakly conserved consensus sequences: the 5′ splice sites, 3′ splice site, polypyrimidine tract (PPT) and branch point sequence (BPS). These elements are necessary but often insufficient to define exon–intron boundaries. Auxiliary *cis*-elements that activate or repress splicing, known as exonic and intronic splicing enhancers or silencers (ESEs, ISEs and ESSs, ISSs), allow the genuine splice sites to be correctly recognized among pseudo-splice sites with similar signal sequences. These *cis*-elements have been identified through mutations that alter splicing, through computational comparisons and through the selection of sequences that activate splicing or bind to splicing regulatory proteins, most notably a family of serine–arginine (SR) proteins (5–9). Disease-causing mutations that affect splicing have contributed to our understanding of splicing signal sequences; however, the influence upon splicing of naturally occurring DNA polymorphisms has not been systematically studied.

Genes encoding histocompatibility leukocyte antigens (HLA) are among the most polymorphic loci in the human genome. A high, selection-driven DNA variability of both class I and class II genes is instrumental for the development of adaptive immune responses; the absent or diminished expression of the latter, results in immunodeficiency (reviewed in 10). Although the *HLA* genes are the most

*To whom correspondence should be addressed at: University of Southampton School of Medicine, Division of Human Genetics, Duthie Building, MP808, Southampton SO16 6YD, UK. Tel: +44 2380796425; Fax: +44 2380794264; Email: igvo@soton.ac.uk

important genetic risk factors for the development of a large number of autoimmune disorders, including multiple sclerosis, rheumatoid arthritis, juvenile diabetes and narcolepsy, it is not known how their AS is controlled.

Alternatively spliced *HLA* isoforms that lack exons encoding the transmembrane domains enter the endocytic compartment, are released by cells and can be detected as soluble molecules (sHLA) in body fluids (11,12). sHLAs have been shown to induce apoptosis of alloreactive cells *in vitro* and modulate immune responses (13,14). Their serum levels are increased in inflammatory and autoimmune diseases and correlate positively with disease activity and autoantibody titres (12). In addition, sHLA levels in normal individuals have been linked to *HLA* haplotypes, with the highest sHLA class II concentration in individuals carrying DR3 and DR4 alleles (15), but mechanisms underlying their haplotype-dependent expression have not been elucidated.

HLA-DQB1 shows variable inclusion of two exons in mature transcripts. Exon 5, which encodes the cytoplasmic portion of DQβ, is incorporated into mRNA only on alleles that have a IVS4−1A→G substitution at the splice acceptor site. Although both alternatively spliced isoforms restrict antigens, transfectants lacking exon 5 were somewhat less efficient at comparable expression levels than those containing this exon (16). Exon 4, which codes for the transmembrane domain of DQβ, is differentially included in mRNA in cell lines carrying particular homozygous *DQB1* genotypes (11), but the molecular mechanism for the allele-dependent production of sDQβ chains is not understood.

Here we show that BPS and PPT variants in intron 3 are major determinants of the allele-specific production of *DQB1* transcripts lacking exon 4 (Δ4). Relative levels of *DQB1* isoforms were markedly influenced by two single nucleotide polymorphisms (SNPs) in BPS, implicating naturally occurring BPS haplotypes in AS. Exon inclusion levels of BPS haplotypes correlated with *in vitro* binding of splicing factor 1 (SF1) to BPS RNAs. In addition, we also demonstrate a biased distribution of disease-causing point mutations in human BPS (hBPS) and propose that impaired SF1−BPS interactions early during spliceosome contribute to splicing defects in these cases. Moreover, we show that AS of this multi-allelic reporter is controlled by at least seven SR proteins and at least three heterogenous nuclear ribonucleoproteins (hnRNPs) and provide a mechanistic explanation for their opposite effects on splice site selection. Finally, we show how subphysiological temperatures alter the allele-specific expression of *DQB1* isoforms and propose that the observed changes in exon inclusion may account for temperature-sensitive course of autoimmune disorders associated with high Δ4.

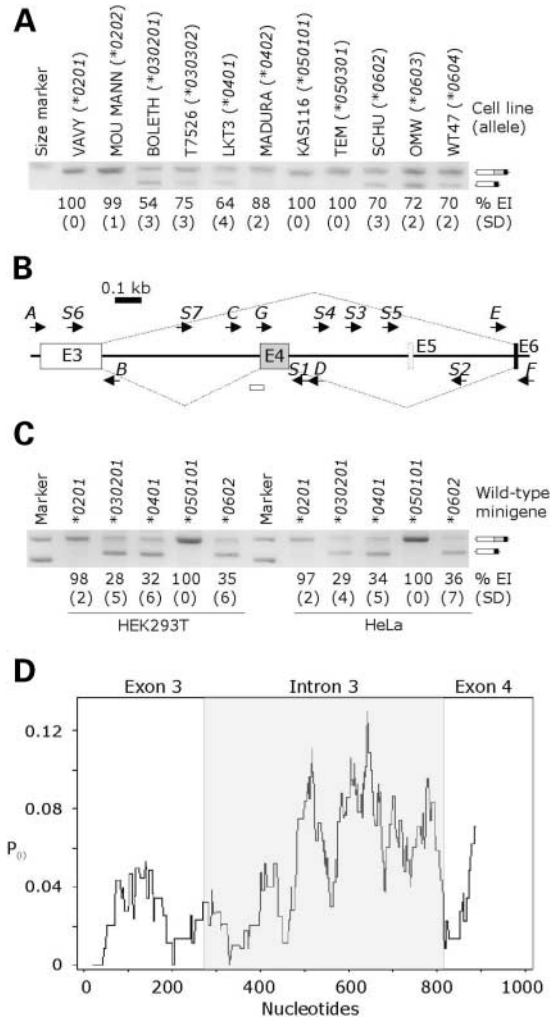

Figure 1. Alternative splicing of *DQB1* exon 4 is determined by *cis*-regulatory elements. (**A**) Differential exon 4 skipping levels in homozygous cell lines. RT−PCR was with upper (T31−32, T37, T41) and lower (T28−29, T40) haplotype-specific primers. Exon inclusion (EI) levels were measured from duplicate cultures. *DQB1* exons 3, 4 and 6 are shown as white, grey and black boxes, respectively. (**B**) Allele-specific minigenes. Exons (boxes, E) and introns (lines) are to scale; scale units are kilobases (kb). All primers (arrows) are shown in Supplementary Material, Table S1. An open rectangle at the 3′ splice site of intron 3 indicates a multiple sequence alignment shown in Fig. 2A. (**C**) Transient Δ4 levels following transfection of wild-type minigenes into HEK293T and HeLa cells. Amplification was with vector primers PL3 and PL4. Mean and SD were derived from three duplicate transfections. (**D**) The nucleotide diversity ($P_{(i)}$, 93) in *DQB1* intron 3. Nucleotide sequences of introns 3−5 are available from Genbank (Accession nos AY553273−AY553292).

RESULTS

Identification of *cis*-acting elements that control differential exon 4 inclusion

To determine factors responsible for the allele-specific production of Δ4, we first characterized *DQB1* exon 4 splicing in a panel of lymphoblastoid cell lines homozygous at *DQB1*. We confirmed the absence of Δ4 in cell lines carrying *DQB1*02* and **05* alleles (Fig. 1A) that do not secrete sDQβ (11). We next constructed minigenes derived from Δ4-producing and non-producing lineages (Fig. 1B) and determined exon inclusion levels after transient transfection of the wild-type constructs into HeLa and 293T cells (Fig. 1C). The allele-specific minigenes retained the signals required for differential exon inclusion in both cell types, indicating that Δ4 variability is determined by polymorphisms within the constructs. To define these *cis*-elements, we sequenced flanking introns in homozygous cell lines carrying 10 *DQB1* alleles that are common in the population. We found

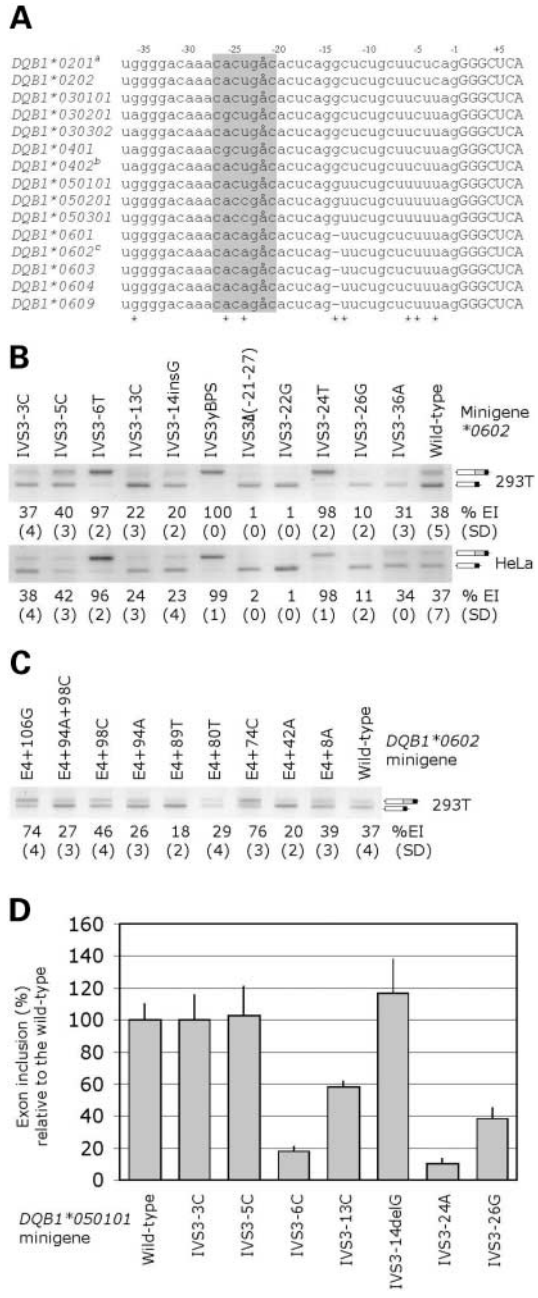

Figure 2. IVS3-24/IVS3-26 in BPS and IVS3-6 in PPT are key SNPs for differential exon 4 inclusion. (**A**) Multiple alignments of *DQB1* pre-mRNAs at the 3′ splice site of intron 3. Intronic sequence is in lower case, exonic sequence is in upper case. Stars denote SNPs, dash indicates a single nucleotide deletion. BPS is shaded and BP adenosine is shown as å. The designation of *DQB1* alleles is according to the World Health Organization Nomenclature Committee (102). The following alleles (cell lines) were sequenced: *DQB1*0201* [cell line VAVY, the 10th International *HLA* Workshop number (IHW) 9023], *DQB1*0202* (MOU MANN, IHW9050), *DQB1*030101* (LUY, IHW9070; TUBO, IHW9045; AMALA, IHW9064), *DQB1*030201* (BOLETH, IHW9031; YAR, IHW9026; PE117, IHW9028; WT51, IHW9029; BM14, IHW9033; SAVC, IHW9034; MT14B, IHW9098), *DQB1*030302* (T7526, IHW9076), *DQB1*0401* (LKT3, IHW9107), *DQB1*0402* (MADURA, IHW9069), *DQB1*050101* (KAS116, IHW9003), *DQB1*050201* (AZH, IHW9293); *DQB1*050301* (WT52, IHW9306; TEM, IHW9057; KOSE, IHW9056), *DQB1*0601* (E4181324, IHW9011), *DQB1*0602* (SCHU; IHW9013), *DQB1*0603* (OMW; IHW9058), *DQB1*0604* (WT47, IHW9063; KOSE, IHW9056), *DQB1*0609* (HO301, IHW9055). All the cell lines were homozygous at *DQB1*, except for KOSE. [a,b,c]The sequences were identical to the GenBank entries U92032, Z80898 and AL662789, respectively. Alignments of exon 4 sequences are available from the IMGT/HLA database (102). (**B**) Exon inclusion levels of minigenes carrying a Δ4-producing allele (**0602*) and mutated in intron 3 SNPs and in BPS. yBPS is the yeast BPS consensus sequence. (**C**) Exon inclusion levels following mutations in exon 4 SNPs. (**D**) Exon inclusion of minigenes carrying a non-skipping allele and mutated in the intronic SNPs.

a total of 65 polymorphisms in intron 3 and 112 polymorphisms in intron 4/5, excluding the sites with alignment gaps. The nucleotide variability was notably higher in the 3′ part of the intron than in the 5′ part (Fig. 1D), suggesting intron homogenization in the region adjacent to β2-encoding exon 3.

To identify SNPs that determine differential splicing, a minigene producing a mixture of exon inclusion and skipping (**0602*) was individually mutated in eight exonic and 10 flanking intronic SNPs revealed by multiple sequence alignments (Fig. 2A and B). IVS3-24A→T and IVS3-6C→T mutations strikingly reduced Δ4, whereas the remaining intronic SNPs, including those close to the 3′ (Fig. 2B) and 5′ (IVS4+14A/G, IVS4+33G/C, data not shown) splice sites, resulted in minor or no Δ4 alterations. Mutations in most exonic SNPs altered Δ4 levels, but to a smaller extent than IVS3-24 or IVS3-6 (Fig. 2C). To confirm the importance of SNPs that influence Δ4, we introduced mutations in the same positions into a minigene that contained IVS3-24T/-6T and produced virtually no Δ4 (**050101*). Measurement of low Δ4 levels with a sensitive real-time RT–PCR assay (Supplementary Material, Fig. S1) showed that maximum exon skipping was conferred by IVS3-24T→A followed by IVS3-6T→C (Fig. 2D). In addition, IVS-26A→G substantially reduced exon 4 inclusion of both **0602* and **050101* minigenes (Fig. 2B and D), indicating that IVS3-26, -24 and -6 were key SNPs responsible for differential splicing of both *DQB1* alleles.

Exon 4 inclusion is determined by BPS haplotypes

The SNPs IVS3-24 and -26 were within the best match to the hBPS consensus (YNYURAY) near the 3′ splice site (Fig. 2A). A pre-mRNA substrate consisting of *DQB1* exons 3 and 4 and complete intron 3 was spliced *in vitro* (Fig. 3A) and the BP was mapped by reverse transcription across the 5′–2′ bond of the BP, followed by inverse PCR (17) and sequencing (Fig. 3B). The identification of A-22 as the BP was further supported by complete exon skipping caused by mutation −22A→G and deletion IVS3Δ-21–27, whereas the replacement of the same heptamer with the yeast consensus UACUAAC (yBPS), which is the preferred mammalian BPS (18), dramatically improved exon inclusion (Fig. 2B). Together, these results indicated that IVS3-21–27, which contains two SNPs, is the BPS of exon 4.

Next, we systematically examined how each of the six BPS haplotypes controls exon inclusion (Fig. 3C). Adenine to uracil substitutions at position −2 relative to BP (BP-2; IVS-24) had the strongest effect. The $A_{-26}T_{-24}$ haplotype gave the lowest Δ4 levels, with a repressive hierarchy $A_{-26}T_{-24} > G_{-26}T_{-24} > A_{-26}C_{-24} > A_{-26}A_{-24} > G_{-26}C_{-24} > G_{-26}A_{-24}$ on Δ4. Accordingly, homozygous cell lines

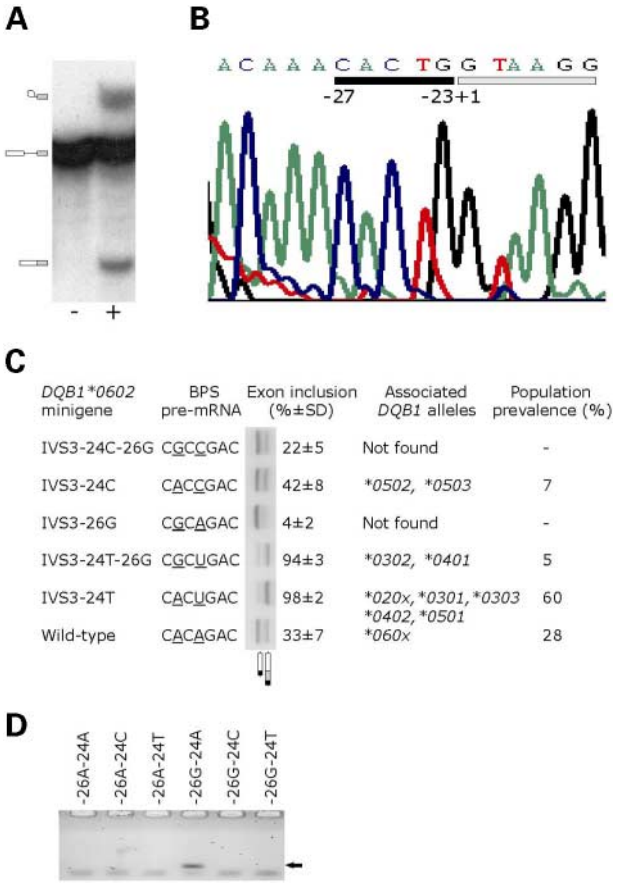

Figure 3. Alternative splicing of *DQB1* exon 4 is controlled by BPS haplotypes. (**A**) *In vitro* splicing of *DQB1*050101* pre-mRNA in the presence (+) and absence (−) of nuclear extracts. The remaining products of the splicing reaction are not shown. (**B**) *DQB1* IVS3 branchpoint adenosine maps to position −22. The black line indicates the 5′ end of IVS3 BPS and the grey line indicates the 5′ end of the intron. BP at A-22 itself was not used as a template during the RT reaction. (**C**) Exon inclusion levels of the *DQB1* constructs mutated in BPS. Mean and SD were derived from three duplicate transfections. The population prevalence of BPS haplotypes was estimated from published frequencies of *DQB1* alleles in white Americans (103), assuming absolute linkage disequilibrium between IVS3-24 and exon 2 specificities. (**D**) Specificity of ARMS–PCR to detect BPS GA-haplotypes. Arrow indicates a GA-containing product amplified using allele-specific PCR with plasmid templates carrying the indicated BPS haplotypes.

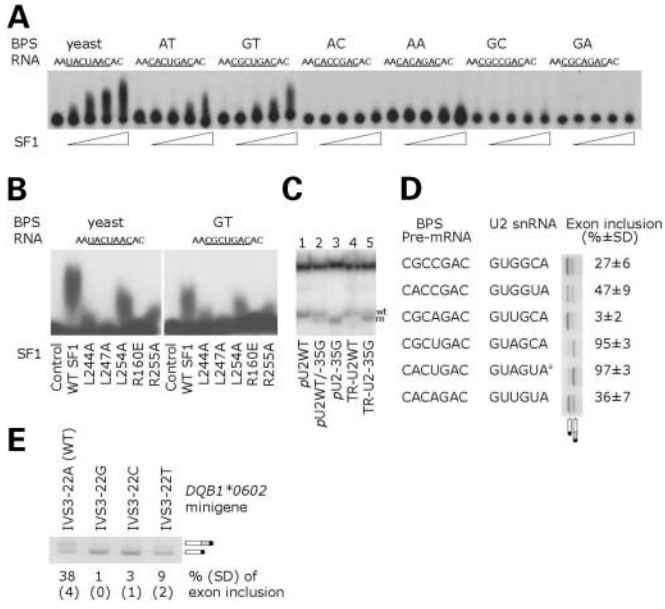

Figure 4. Impaired SF1 interactions with BPS RNAs. (**A**) Wild-type SF1. Recombinant SF1 (0.68, 1.36, 2.72 and 5.44 μg) was incubated with short RNAs representing BPS haplotypes as described (21). (**B**) Mutated SF1. Mutations were introduced in residues that interact with BP-2 (L244, L247, R160) and BP-4 (R255 and possibly L254; 21). (**C**) Expression of mutated U2 snRNA in transfected cells. A single-strand conformation polymorphism gel with DNAs extracted from the wild-type and mutated U2 snRNA plasmids (lanes 1 and 3), their 1:1 mixture (lane 2) and with cDNAs prepared from HEK293T cells transfected with wild-type (lane 4) and mutated (lane 5) U2 snRNA. (**D**) Exon inclusion levels of the *DQB1*0602* minigenes mutated in BPS SNPs and co-transfected with wild-type (indicated with superscript a) and mutated U2 snRNA plasmids. Exon inclusion in controls transfected with the same amount of vector DNA (data not shown) was not significantly different from that in untransfected cells (cf. Fig. 3C). (**E**) Exon inclusion levels of minigenes mutated at BP.

carrying the $A_{-26}T_{-24}$ haplotype produced less Δ4 than cells with other BPS haplotypes (Fig. 1A and C). However, BPS variation alone could not explain high Δ4 observed in $G_{-26}T_{-24}$-carrying cell lines, implicating IVS3-13 and other intronic SNPs in differential exon skipping, as exon 4 is identical on the *DQB1*03/*04* alleles.

BPS haplotypes associated with the lowest exon inclusion are absent in the population

Initial sequencing of *DQB1* alleles in homozygous cells did not identify the $G_{-26}A_{-24}$ and $G_{-26}C_{-24}$ haplotypes that yielded the highest Δ4 (Fig. 3C). We therefore sequenced BPS of additional *DQB1* alleles associated with IVS3-26G or -24A/C. Without exception, homozygous cell lines carrying IVS3-26G and -24A/-24C had -24T and -26A, respectively (Fig. 2A). In addition, genotyping of over 1600 Caucasian chromosomes using a haplotype-specific (Fig. 3D) amplification refractory mutation system (ARMS)–PCR (19) did not reveal these haplotypes. These observations prompt the hypothesis that these intronic sequences may have been subject to negative selection through diminished expression of natural transcripts, leading to immunodeficiency and predisposition to infections.

Impaired interactions of SF1 with BPS RNAs that produce high Δ4

During spliceosome assembly the BPS is recognized first by splicing factor 1 (SF1; branchpoint binding protein) in the early (E) complex (20,21), and later by base-pairing with U2 snRNA in the A complex (22–24). The influence of BPS haplotypes on exon 4 splicing could be mediated by the interactions of either or both of these *trans*-acting factors with the BPS.

To test if SF1 binding to BPS haplotypes correlates with exon inclusion, we examined the interaction between small synthetic BPS-containing RNAs with wild-type and mutated

Table 1. Single-nucleotide substitutions in hBPS that resulted in recognizable phenotypes

Gene	Y_{-5}	N_{-4}	Y_{-3}	U_{-2}	R_{-1}	A_0	Y_{+1}	Splicing abnormality	Reference	Intron length (nt)[a]
L1CAM	A	T	C	C	A	A→C[b]	G	Activation of a cryptic 3′ splice site	Rosenthal *et al.* (104)	829
LIPC	C	C	C	C	A	A→G	T	Activation of cryptic 3′ splice sites	Brand *et al.* (105)	$>10^5$
FBN2	T	T	G	C	A	A→G	T	Exon skipping	Putnam *et al.* (106)	410
HEXB	T	T	G	C	A	A→G	T	Activation of a cryptic 3′ splice site	Fujimaru *et al.* (107)	433
XPC[c]	T	A	C	T	G	A→G	T	Exon skipping	Khan *et al.* (108)	2057
HEMB	C	G	T	T	A	A→G	T	Not determined	Ketterling *et al.* (109)	188
TSC2	G	C	G	T	G	A→G	C	Activation of cryptic 3′ splice site and intron retention	Mayer *et al.* (110)	106
LCAT[c]	C	C	C	T→C	G	A	C	Complete intron retention	Kuivenhoven *et al.* (31)	83
ITGB4[c]	G	G	C	T→A	C	A	C	Intron retention and activation of a cryptic 3′ splice site	Chavanas *et al.* (30)	88
TH[c]	G	G	C	T→A	G	A	T	Exon skipping and cryptic 3′ splice site activation	Janssen *et al.* (111)	238
COL5A1	G	A	C	T→G	G	A	T	Exon skipping	Burrows *et al.* (32)	5875
DQB1[c]	C	G/A	C	T/A/C[b]	G	A	C	Variable exon skipping	This study	509–516

[a]Intron length was determined from www.ensembl.org, except for *DQB1*.
[b]Arrows and slashes indicate naturally occurring mutations and polymorphisms, respectively.
[c]BP was determined experimentally as described in each reference. Minigene constructs derived from the remaining genes showed impaired splicing following mutations of the indicated BP adenosine (J. Kralovicova *et al.*, manuscript in preparation).

SF1 in gel mobility shift assays. BP-2U-carrying RNAs (haplotypes $A_{-26}T_{-24}$ and $G_{-26}T_{-24}$, Fig. 4A) showed significant gel shifts, but less than the yBPS control, consistent with a higher exon inclusion of the yBPS- than IVS3-24T-containing minigenes (Figs 2B and 3C). In contrast, SF1 binding to BP-2C or BP-2A RNAs (haplotypes $A_{-26}C_{-24}$, $A_{-26}A_{-24}$, $G_{-26}C_{-24}$, $G_{-26}A_{-24}$) was weak or absent even at the highest SF1 concentration. SF1 mutated in residues that interact with BP-2 or BP-4 (21) had impaired binding to yBPS and a BP-2U BPS (Fig. 4B). Exon inclusion levels of BPS mutants correlated with SF1 binding ($r = 0.88$, $P = 0.02$, Supplementary Material, Fig. S2), suggesting that exon 4 splicing was determined to a large extent by the strength of specific SF1–BPS interactions *in vitro*.

Next, we tested if U2 snRNA–BPS base-pairing can alter exon inclusion. We cotransfected *DQB1*0602* reporters with plasmids expressing U2 snRNA and carrying complementary mutations in the BPS-interacting region. Surprisingly, despite high expression of mutated U2 snRNA (Fig. 4C) and repeated experiments with several vectors, we observed no significant decrease of transient Δ4 (Fig. 4D).

Predominance of hBPS point mutations in BP and BP-2

Comparison of previously published disease-associated single-base substitutions in hBPS showed that these alterations occurred only in BP and BP-2 (Table 1). Their biased distribution across BPS ($p = 10^{-6}$, binomial test for both BP and BP-2) suggests that point mutations in the remaining BPS positions give rise to no or only minor splicing defects and do not produce recognizable phenotypes. Since BP-2U mutations markedly impair SF1–BPS interactions (Fig. 4A) (20,25) and BPSs in naturally occurring BP-2U substitutions (Table 1) were similar to those in *DQB1* IVS3 (Fig. 2A), the associated splicing abnormalities of hBP-2 mutations are likely to arise through weakened SF1 binding. Because BP mutations efficiently reduce these interactions as well (20,25), the majority, if not all, single nucleotide substitutions in hBPS may affect splicing via impaired SF1–BPS binding.

A predominance of A→G transitions in BP (Table 1) suggested that hBP mutated to pyrimidines, which can be used for productive lariat formation (26), generate more natural transcripts than BP-A→G transitions, consistent with a more severe reduction of the first and second splicing steps *in vitro* observed for guanine-containing substrates (27). To test this *in vivo*, we examined *DQB1* isoforms after transfection of the **0602* minigenes mutated at BP into 293T cells. The transition mutant (BP-A→G) gave the highest expression of transient Δ4, whereas both transversions yielded lower amounts (Fig. 4E). However, because BP transversions and transitions in very short introns may generate similar levels of exon skipping (28), their phenotypic outcome is likely to be influenced by intronic sequences. This is further supported by comparison of the splicing phenotype of human BPS alterations with intronic length, suggesting that mutations in very short introns, particularly those in BP-2, tend to result in intron retention, whereas those in longer introns usually skip downstream exons (Table 1).

Phenotypic consequences of BPS substitutions are influenced by the intronic context

To further test these predictions, we examined minigenes with truncated introns (Fig. 5A). When compared with the wild-type minigene, the truncated **0602* minigene showed negligible exon 4 skipping and substantial intron 3 retention (Fig. 5A, lane 2). In contrast, a chimeric construct, in which exon 4 and flanking intronic sequences were derived from the Δ4-lacking allele (**050101*, Fig. 1C), produced transcripts that retained only minimal amounts of intron 3 (Fig. 5A, lane 1). Since the two minigenes diverged in only four SNPs, we mutated each position in the former construct and examined the splicing pattern. IVS3-24A→T showed the most prominent reduction of intron 3 inclusion, whereas mutations in the

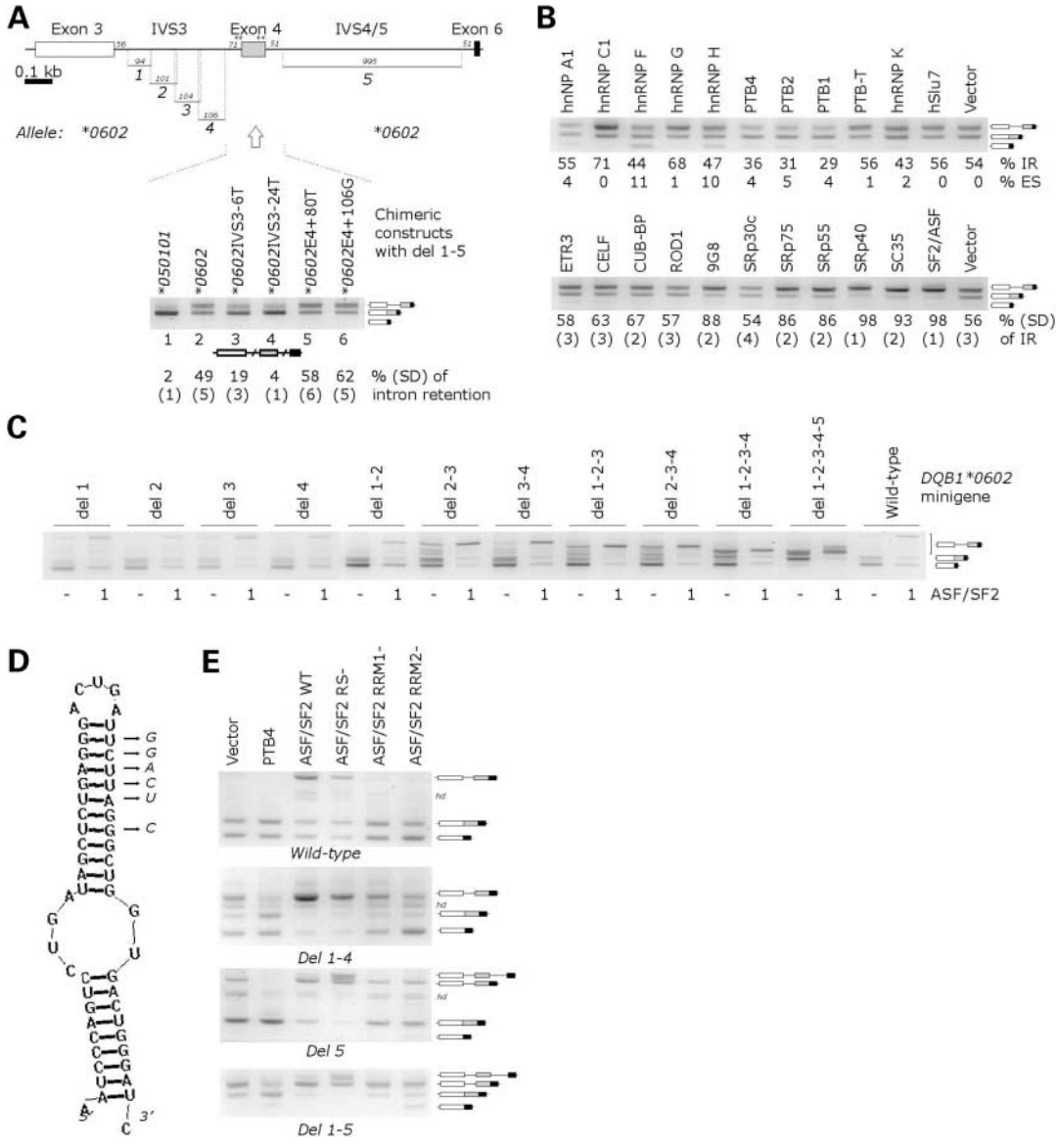

Figure 5. Opposite effects of hnRNPs and SR proteins on *DQB1* splicing. (**A**) Truncated and chimeric *DQB1*0602* constructs. The size of deleted segments named 1–5 is in nucleotides. Stars denote SNPs (upper panel). Arrow indicates a replacement of exon 4 and flanking intronic sequences with segments specified in the lower panel. (**B**) Opposite effects of SR proteins and hnRNPs I, H and F on intron 3 retention. One microgram of plasmid DNA expressing the indicated factors was co-transfected with 0.5 μg of the *DQB1*0602* reporter lacking segments 1–5. ES, exon skipping; IR, intron retention. (**C**) The influence of intron deletions on the ASF/SF2-induced splicing repression. Reporter constructs shown in (A) were co-transfected with plasmids expressing the wild-type ASF/SF2 (the amount is in micrograms). (**D**) Mutations introduced in the predicted stem structure in segment 2–3. (**E**) PTB- and ASF/SF2-mediated alterations of exon skipping and intron retention in the wild-type and truncated **0602* constructs. Equimolar amounts of constructs lacking RRM1, RRM2 and RS-domains were cotransfected with 0.5 μg of *DQB1* reporter constructs shown below each panel; *hd*, heteroduplexes.

remaining SNPs contributed much less to retention (Fig. 5A), further supporting the importance of BP-2 uracil in intron 3 removal and exon 4 skipping. Diverse splicing and clinical outcome of BP-2T substitutions is thus determined by intronic sequences, most likely involving intron–bridging interactions mediated by SF1, consistent with the central role of this protein in yeast cross-intron contacts (29) and with the retention of short (30,31), but not long introns (Fig. 1A and C) (32) in disease-causing BP-2T substitutions.

Opposite effects of SR proteins and hnRNPs on *DQB1* splicing

To examine the influence of protein factors known to regulate splicing on both exon 4 inclusion and intron 3 retention, we individually cotransfected *DQB1* reporters with plasmids expressing SR proteins, a subset of hnRNPs and several members of the CUG-BP and ETR-3 like factor (CELF) protein family. Surprisingly, transcripts retaining intron 3 were

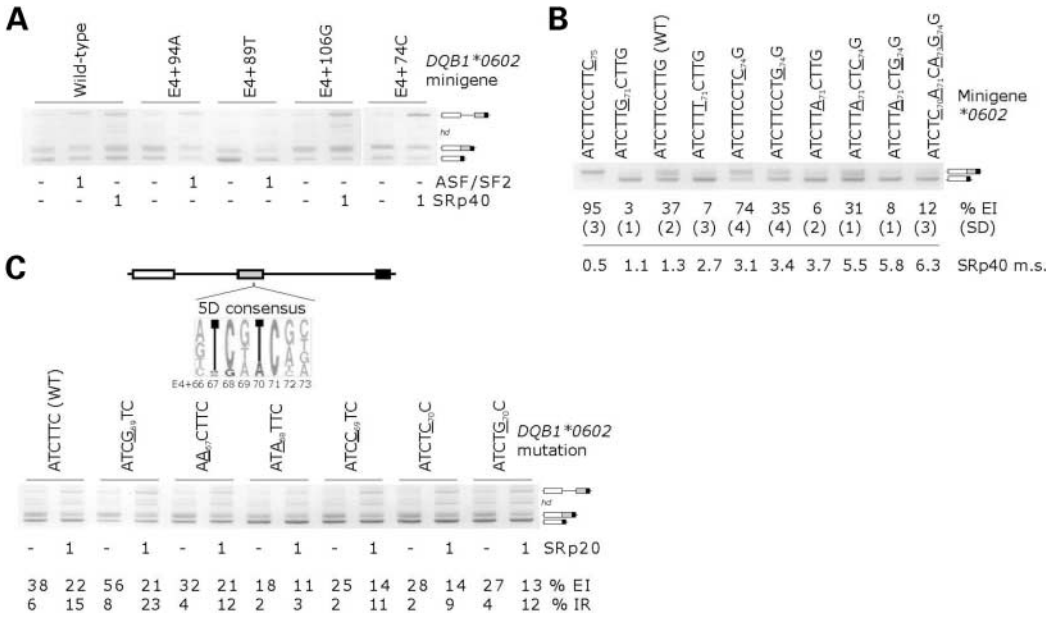

Figure 6. Identification of splicing regulatory sequences in *DQB1* exon 4. (**A**) Co-transfections of *DQB1* reporters mutated in positions predicted by functional SELEX (34) with SRp40 and wild-type ASF/SF2. The amounts of plasmid DNA are in micrograms. (**B**) Influence of mutations in a putative SELEX binding site for SRp40 on exon inclusion. Mutations are underlined and their positions in exon 4 are indicated with a subscript. Matrix scores (m.s.) for SRp40 were derived as described (34). Mutation 75C introduced a putative binding site for SRp55 (m.s. 3.01), while mutations maximizing the SRp40 score introduced putative binding sites for ASF/SF2 (m.s. 5.2). (**C**) Influence of mutations in the 5D consensus (8) on exon inclusion (EI) in cells co-transfected with SRp20. The amount of SRp20 is in micrograms. Pictogram of ESE 5D is reproduced with permission from *Science* (8).

increased in cells overexpressing SR proteins in both the truncated (Fig. 5B) and the full-length (data not shown) minigenes. In contrast, transcripts that retained intron 3 were diminished in cells overexpressing hnRNP I (polypyrimidine binding protein, PTB), F and H. The splicing pattern induced by alternatively spliced PTB isoforms was similar, except for PTB-T, which lacks exons 3–10 and RNA recognition motifs (RRM) 1 and 2 (33). Among constructs with serial deletions of ~100 bp segments in intron 3 (Fig. 5A), deletion of segment 3 generated maximum ASF/SF2-induced intron retention (Fig. 5C). This segment was predicted to form significant RNA secondary structure, but multiple substitutions designed to remove the most prominent stem–loop (Fig. 5D) had no effect on intron retention or exon skipping both in the wild-type and shortened **0602* minigenes (data not shown).

Figure 5E shows *DQB1* transcripts after cotransfection of the wild-type and truncated reporters with plasmids expressing ASF/SF2 and PTB. Truncations in either intron markedly elevated levels of intron 3-retaining transcripts. In contrast, exon skipping was promoted by removal of intron 3 sequences, but was eliminated by truncation of the second intron. The opposite effects of ASF/SF2 and PTB were dose-dependent and cotransfection of a constant PTB amount with increasing amounts of ASF/SF2 and vice versa antagonized the effects of either protein (data not shown). Importantly, ASF/SF2 lacking RRM1 or RRM2 lost the intron-retaining activity (Fig. 5E). The absence of RRM2, but not RRM1, promoted exon skipping of reporters truncated in intron 4/5. Interestingly, a lack of RS domain did not eliminate repression of intron 3 splicing, but enhanced retention of intron 4/5. RNA binding was thus essential for the PTB-mediated activation and the ASF/SF2-induced repression of intron 3 splicing, whereas the RS domain was dispensable for the ASF/SF2-mediated retention of weakly spliced intron 3, but not strong intron 4/5.

Exon 4 inclusion is controlled by ESE 5D

Most constructs mutated in exon 4 SNPs reproducibly increased (E4+74, E4+98, E+106) or reduced (E4+89, E4+42, E4+80, E+94) exon inclusion. Because Δ4 expression was most influenced by E4+74 (Fig. 2C), we examined this SNP in more detail. A functional SELEX tool (34) that may facilitate identification of ESEs responsive to a subset of human SR proteins predicted that E4+74T/C may create a putative binding site for SRp40 (TTCCT$\underline{C}_{74}$G). In contrast, E4+89C/T and E4+94A/G may reduce the ASF/SF2 interaction and E4+106G/A may impair SRp40 binding (arrows in Supplementary Material, Fig. S3). In cotransfection experiments with plasmids expressing the two SR proteins, both the wild-type and mutated clones showed an increase of intron 3 retention, but exon inclusion levels of mutated clones were similar to the wild-type minigenes (Fig. 6A). Examination of a series of constructs mutated in the flanking sequence to maximize and minimize SRp40 matrix scores revealed considerable alterations of exon inclusion, indicative of the presence of strong ESE/ESSs, but there was no significant correlation between the matrix scores and exon inclusion ($P = 0.2$, Fig. 6B).

Interestingly, an ESE derived by the relative enhancer and silencer classification by unanimous enrichment (RESCUE) and termed 5D (8) was present in minigenes that showed high exon inclusion, but was absent in constructs with high

Δ4. Mutations in the conserved 5D positions (nucleotides 66–71) to nucleotides that were rare or absent in the 5D consensus significantly increased Δ4. The highest increase was observed for mutations of the last position (E4+71C), which is invariant in this hexamer. In contrast, mutation 69T→G, which improves the match to 5D, lowered Δ4 (Fig. 6C). As with other SR proteins (Fig. 5B), cotransfection of mutated constructs with SRp20, which may bind 5D (8,35,36), increased intron 3 retention (Fig. 6C); however, the SRp20-induced increase of Δ4 was observed for each 5D mutation.

DQB1 exon 4 skipping is increased at subphysiological temperatures

The spliceosomes assembled on different splice sites are not identical and physiological conditions may therefore differentially affect the splicing pattern of various messages. As temperature has ambivalent effects on splicing efficiency and exon inclusion (37) and BPS mutations are often associated with temperature-sensitive phenotypes in lower organisms (38), we measured Δ4 using real-time PCR in homozygous cell lines maintained at various temperatures prior to RNA extraction. Cells kept at 30°C had the highest Δ4 expression relative to natural transcripts (Fig. 7A). The Δ4 levels in cells maintained at 5°C were similar to those observed for physiological temperatures, whereas cells kept at 20°C had intermediate levels. Cell lines exhibiting the highest Δ4 at 37°C had the peak Δ4 levels also at non-physiological temperatures, indicating that the differential exon skipping is maintained under stress conditions. Resuscitation of cells kept at lower temperatures for additional 24 h at 37°C prior to RNA extraction normalized Δ4 levels (data not shown). Exon skipping was also significantly increased for transiently expressed haplotype-specific minigenes in cells maintained at 30°C (Fig. 7B). Thus, subphysiological temperatures (20 and 30°C) enhanced both the endogenous and exogenous Δ4 expression across the cell lines.

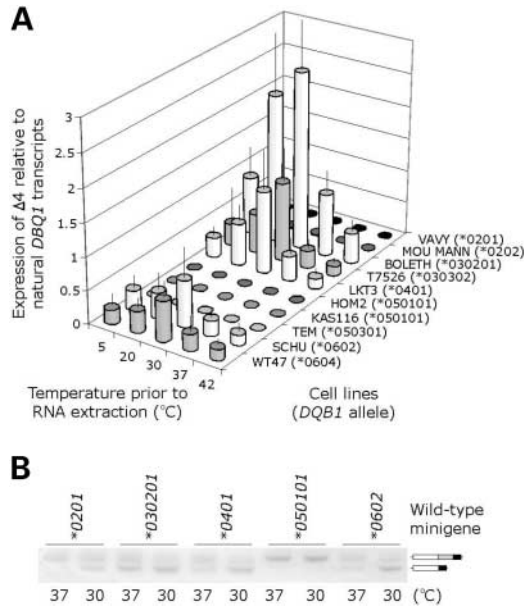

Figure 7. Temperature-sensitivity of exon 4 inclusion. (**A**) Endogenous Δ4 expression in lymphoblastoid cell lines homozygous at *DQB1*. Cells were grown at the indicated temperatures for 16 h prior to RNA extraction. Exon skipping was measured using real-time RT–PCR (Supplementary Material, Fig. S1). Bars indicate SDs from two experiments. (**B**) Transient Δ4 expression of allele-specific minigenes in HEK293 cells maintained at the indicated temperatures for 16 h prior to cell lysis. RT–PCR was with vector-specific primers.

DISCUSSION

We have shown the first examples of naturally occurring BPS haplotypes [*DQB1* IVS3-26(A/G), -24(A/C/T)] that influence gene expression, providing the clearest illustration to date of haplotype control of AS. Because sDQβ are present in culture supernatants of Δ4-positive, but not Δ4-negative cell lines (11), we propose that their haplotype-specific secretion is largely determined by IVS3 PPT and BPS variants through differential recognition of BPS by the KH-QUA2 domain of SF1. Our negative U2 snRNA complementation data (Fig. 4D) appear to suggest that altered U2 snRNA–BPS base-pairing, which is obligatory in yeast with virtually invariant BPS, may not necessarily modify splicing of some human introns. This notion seems to be supported by a similar observation in human β-globin (24) and by several examples of 'U1 snRNA-independent' pre-mRNA substrates (39–44). In addition, if U2 snRNA–BPS interactions substantially contributed to altered splicing caused by phenotype-associated BPS point mutations (Table 1), one would expect to see a more uniform distribution of these changes across hBPS. Interestingly, the two positions (BP and BP-2) have been shown to be the most effective in reducing SF1–BPS interaction (20,25). However, without a positive control for the U2 snRNA complementation, we could not rule out inefficient entry of transient U2 snRNA into U2 snRNPs, although RNA preparations were not enriched for snRNPs in previous studies (23,24). Moreover, we could not completely exclude cryptic BP usage in some constructs, including those with PPT SNPs, which may alter interaction of the 65 kDa subunit of U2 auxiliary factor ($U2AF^{65}$) with the pre-mRNA and cooperative interaction between SF1 and $U2AF^{65}$ that facilitates BPS recognition (45,46). However, our BP mapping showed only a single inverse PCR product, a minigene with deleted BPS completely repressed splicing (Fig. 2B) and constructs with mutated BP-22A showed differential exon inclusion (Fig. 5E), consistent with differential recognition of these BPS substitutions *in vitro* (27). Finally, it is also possible that additional BPS-interacting factors, including SR proteins (47–49), might play a role in AS of *DQB1* and in splicing defects caused by human BPS substitutions (Table 1).

Membrane proteins are commonly regulated by mechanisms that release their soluble forms from its membrane anchorage (50). Transmembrane domains of a large number of human and mouse proteins are removed by AS more frequently than expected and often in a tissue-specific manner (51,52), suggesting that these events are biologically significant. Membrane domains interact during the assembly of class II heterodimers (53), influence the organization of endosomal compartments (54), facilitate DM colocalization and regulate peptide exchange (55). Although Δ4 are productive

transcripts (11) that may modulate immune responses (13,14) and bind similar peptides as membrane-bound molecules (56), physiological significance of sDQ and their putative role in susceptibility to autoimmune disease have been obscure. Interestingly, the highest concentration of sHLA class II molecules has been observed in individuals carrying DR3 and DR4 alleles (15), which are in linkage disequilibrium with the Δ4-producing *DQB1*0302* and **0401* (Figs 1A and 2A). In addition, higher levels of sHLA class I molecules were found for *HLA-A24* (12), which is linked to *DQB1*0302* and **0401*. Future studies should therefore confirm whether the high Δ4 production is indeed reflected in increased levels of sDQβ in body fluids.

Excessive amounts of Δ4 observed for some alleles (Fig. 7) provoke the hypothesis that the observed temperature-induced mRNA alterations could influence surface expression of DQ heterodimers, interactions of the peptide–MHC complexes with T-cell receptor (TCR) and, ultimately, the course of (or susceptibility to) autoimmune diseases. Most notably, joint pain in patients with rheumatoid arthritis (RA), which is associated with Δ4-producing *DQB1*030x* and **040x* alleles, has long been linked to low temperatures (57). It can also be speculated that a well-known latitude gradient in the population prevalence of autoimmune disorders, sometimes attributed to pathogen distribution, ultraviolet radiation or other environmental or genetic effects, may be influenced by the temperature gradient. Since temperature-dependency of RA and other autoimmune disorders, such as systemic lupus, Behcet disease and ankylosing spondylitis (58), may also result from altered temperature-induced association and dissociation rates of peptide–MHC interactions with TCR (59), it will be interesting to examine these possibilities in future studies.

Although subphysiological temperatures enhanced exon 4 skipping of each allele (Fig. 7), they may also elicit exon inclusion (60) or inhibit utilization of the cryptic splice site of other substrates (37). In addition, heat shock has been shown to activate an alternative 5′ splice site (61) and mutation of the last exon nucleotide produced an aberrant transcript not detectable at 31°C, which gradually increased up to 39°C (62). Our results thus further underline the importance of fast processing of clinical samples intended for RNA analysis and temperature control during their transport and storage to minimize the number of incorrectly spliced transcripts and to avoid false positive reports of gene deletions, AS or haplotype-specific mRNA expression.

Alterations of exon inclusion induced by the majority of exonic SNPs and mutations in putative binding sites for splicing regulatory factors support a high abundance of sequences that influence pre-mRNA splicing. Interestingly, the range of exon inclusion levels produced by mutations in natural variants was smaller than for those introduced in putative splicing regulatory regions identified by RESCUE-ESE or SELEX approaches (Figs 2C and 6B). Alterations of exon inclusion following mutations in exonic SNPs were also less variable than in previous studies employing random mutations in coding regions (63,64). For example, 13 of 26 *CFTR* exon 9 mutations lowered exon inclusion under 15% (range 8–98%) (64), whereas none of our constructs mutated in exonic SNPs reduced natural transcripts below this level (range 18–76%, Fig. 2C). This suggests that mutations in the coding region that reduce exon inclusion below an exon- or gene-specific threshold are subject to significant selection at the level of pre-mRNA splicing.

Three of eight tested exonic SNPs (+42, +98 and +106) were located in putative RESCUE-ESE-derived hexamers (8). Mutations in two of them (+106G and +98C), which were close to the 5′ splice site and encoded non-conservative amino acid substitutions, were predicted to remove enhancers, but they significantly reduced Δ4 (Fig. 2C). E4+42A, which was expected to retain ESE 5A/3G (8), had the opposite effect, indicating that the predictive value of putative ESEs to activate splicing is very limited without considering other interactions. In contrast, mutations in ESE 5D located in the middle part of the exon generated predicted exon inclusion levels, including the 68C→A transversion (Fig. 6C), which corresponds to the mutation that had caused exon-skipping in a heterologous context (8). In addition, 5D is likely to contribute to exon definition across species since the AUCUUC consensus is invariant in the mouse (GenBank accession nos K00008, K00111, M60562 and K00116), rat (M15356), pig (AY459305), cow (AY212029), sheep (Z27401), goat (AY464653), cat (AY152833), dog (Y07947) and possum (AY271265).

SR proteins typically promote exon inclusion through binding to exonic ESEs, but they may also repress splicing *in vitro* and *in vivo* (Fig. 5; 65–70). Similarly, PTB has been regarded as a general repressor of splicing (71), but overexpressed PTB can also promote exon inclusion (Fig. 5) (72,73). The observed opposite effects of hnRNPs and SR proteins on splicing of intron 3 and at least some of the previous, seemingly contradictory, results can be reconciled by antagonistic effects of these factors on splice site selection. Although exceptions have been noted (74), ASF/SF2 and other SR proteins generally activate proximal splicing, whereas distal splicing is promoted by hnRNP A1 (65,66,75–78). *DQB1* intron 3 is very weak, with the 5′ splice sites of exon 3 and exon 4 effectively competing for splicing to exon 6. Therefore, the SR-induced intron 3 retention can be explained by a repression of the distal 5′ splice site (i.e. 5′ splice site of exon 3) and/or promotion of the proximal 5′ splice site (i.e. 5′ splice site of exon 4). Similarly, hnRNP-mediated decrease of intron 3-containing transcripts would result from a repression of the proximal splice site and/or promotion of distal splicing. Some factors did not alter isoform ratios (Fig. 5B), but we could not exclude poor transient expression over high endogenous levels, as we did not have antibodies against hnRNP A1, hSlu7 and CELF proteins. However, the negative result with hnRNP A1 (Fig. 5B) resembles a previous observation (79) and may reflect substrate-specific requirements. Despite this drawback, examination of a sizeable panel of hnRNPs and SR proteins, which colocalize in discrete regions of the nucleus (80,81), highlights the opposite effects of the two protein families on the selection of competing splice sites. It also suggests that the network of antagonistic interactions that balance splice site selection may be more extensive and that future studies may substantially add to currently known examples (65,66,75–78,82–85). In addition, the novel multi-allelic splicing reporter system described here will be useful to study interactions across the highly polymorphic splice

acceptor site and to test whether the requirement for U2 snRNA base-pairing with BPS is alleviated by SR proteins, as proposed earlier for U1 snRNP interactions with the 5′ splice site (40,41).

A prominent decrease of Δ4 observed for IVS3-6C→T mutations (Fig. 2B and D) can be explained by improved interactions of PPT with RRM1 of U2AF^{65} or competition with other RNA binding proteins that preferentially recognize polyuridine-rich sequences (86,87). The pentauridine stretch characteristic of the **05* lineage (Fig. 2A) was sufficient to maintain high exon inclusion in the cell line TEM despite weakened BPS of the *DQB1*050301* allele (Figs 1A and 2A). Although uridines are generally the preferred pyrimidine in PPT (88,89), IVS-6U→C substitutions may (Fig. 2B) or may not (90) affect exon inclusion. Differential outcome of T/C substitutions in identical PPT positions relative to the 3′ AG could be due to intron-specific binding registers of U2AF^{65} (87) and/or the presence (90) or absence (Fig. 2A) of purines in the vicinity of the 3′ splice site that weaken PPT recognition. Because RRM1 of U2AF^{65} does not crosslink to a unique set of residues at the 3′ end of PPT (87), prediction of splicing phenotypes caused by point mutations in this region may be more problematic than those in hBPS.

MATERIALS AND METHODS

Multi-allelic splicing reporter system

A 2.2 kb region containing exons 3–6 was amplified using primers A and F (Fig. 1B) with the Expand Long Template PCR System (Roche, USA) and DNA from cell lines homozygous at *DQB1* (Figs 1A and 2A). Primers are shown in Supplementary Material, Table S2. Homozygosity and allele-specificity of template DNA was confirmed with a panel of *HLA* microsatellites and sequence-based typing of *DQB1* and *DRB1* as described (91). Gel-purified DNA was cloned into pCR3.1 (Invitrogen, USA). Mutated minigene constructs were prepared by overlap-extension PCR and validated by sequencing as described (92).

Sequence variability of *DQB1* introns

Introns 3–5 of 10 *DQB1* alleles were sequenced directly using amplicons obtained from homozygous DNA with PCR primers A-S1, A-F or G-F. Sequencing primers S1–S7 are schematically shown in Figure 1B. The nucleotide diversity or the average number of nucleotide differences per site ($P_{(i)}$) (93) was computed using DNA SP software (version 4) (94) with pairwise removal of the alignment gaps (Fig. 1D). The window and step sizes were 30 and 2 nucleotides, respectively.

Expression plasmids and antibodies

Plasmids expressing *ROD1* (95), *PTB1*, *PTB2* (96), *PTB4* (97), *PTB-T* (33), *ETR-3*, *CUG-BP*, *CELF4* (83,98), 9G8, SC35, ASF/SF2, SRp30c, SRp40, SRp55, SRp75 (81), U2 snRNA (24), hSLu7 (99), hnRNP A1, H and K were obtained as a generous gift from G. Ast, D. Black, J. Caceres, T. Cooper, A. Krainer, A. Mayeda, H. Okayama, R. Reed, Z. Ronai, W. Rigby, G. Screaton, C. Smith and A. Weiner. hnRNPs C1, F and G were cloned from a cDNA library using primers shown in Supplementary Material, Table S2. pET-15b-hnRNP H obtained from D. Black was subloned into *Xho*I and *Bam*HI sites of pCR3.1. Mutated ASF/SF2 constructs were obtained from J. Caceres. In addition to previously described construct (24), U2 snRNA was cloned into pCR3.1 and pGEM using primers U2-F/R. All clones were validated by nucleotide sequencing before transfection. The expression of a subset of proteins was confirmed by Western blotting with whole-cell lysates using antibodies against hnRNP I, H, F and K generously provided by D. Black, G. Dreyfus, T. Tamura and C. Smith.

Homozygous cell lines

Cell lines or DNA samples carrying the indicated alleles (Figs 1A and 2A) were purchased from the European Cell Culture Collection (Salisbury, UK) or were a gift from S. Marsh and N. Mayor, Anthony Nolan Research Institute. Cell lines were grown in RPMI1640 with 1% glutamine (v/v) supplemented with 10% fetal calf serum. DNA was extracted as described (92). Total RNA was isolated using RNeasy Mini Kit (Qiagen, Germany). First-strand cDNA was reverse transcribed using oligo(dT)$_{15}$ primers and Moloney murine virus reverse transcriptase (M-MLV H$^-$; Promega, USA) according to manufacturer's recommendations.

Transfections

HeLa and 293T cells were grown under standard conditions in RPMI1640 supplemented with 10% (v/v) fetal calf serum (Gibco BRL, USA). Transient transfections were performed in 6- or 12-well plates using FuGENE 6 (Roche) according to manufacturer's recommendations. Cells were harvested with trypsin 48 h post-transfection.

Detection of *DQB1* isoforms

Total RNA was isolated using RNeasy Mini Kit (Qiagen) or Tri-reagent (Sigma, USA) and treated with DNaseI (Ambion, USA). First-strand cDNA was reverse transcribed using oligo(dT)$_{15}$ primers and Moloney murine virus reverse transcriptase (M-MLV H$^-$; Promega) according to manufacturer's recommendations. Exon inclusion levels were measured on ethidium bromide-stained gels with FluorImager 595 using FluorQuant and Phoretix software (Nonlinear Dynamics Inc., USA). DNA bands were extracted from the gel and sequenced to confirm the identity of each fragment. The number of PCR cycles was 28 or lower to maintain approximately linear relationship between the RNA input and signal. Design of real-time RT–PCR with primers/probes T26–T45 is shown in Supplementary Material, Figure S1. Standard curves were constructed using serial dilutions of DNA extracted from plasmids containing allele-specific inserts cloned using amplicons that were derived from lymphoblastoid cell lines homozygous at *DQB1* (Figs 1A and 2A).

Detection of BPS haplotypes and U2 snRNA expression

A 153 nucleotide fragment containing the BPS-interacting region of U2 snRNA was amplified using single-strand confirmation primers (SSCP)-F/R (Supplementary Material, Table S2). SSCP assays were carried out as described (92). ARMS–PCR (19) for the population screening of BPS associated with the lowest exon inclusion (Fig. 3C and D) was with DNA samples ascertained as described previously (100). Ethics approval was granted by a local committee. The specificity of ARMS–PCR was tested using DNAs isolated from plasmids mutated in BPS (Fig. 3C).

RNA secondary structure predictions

The lowest free energy structures were predicted using enhanced free energy minimization algorithms and plotted (Fig. 5D) with the RNA Structure (version 3.7) (101).

Gel mobility shift assays

Gel shift experiments were carried out as described (21). Ten microlitre reactions contained 32 mM Tris–HCl (pH 8.0), 160 mM NaCl, 8 mM imidazole, 4 mM β-mercaptoethanol, 2.5 μg tRNA, 0.01 pmol of 5′ [^{32}P]-ATP-labelled oligoribonucleotides (Biospring, Germany).

Branchpoint mapping

A two-exon fragment amplified with primers A and D (Fig. 1B) using homozygous DNA was cloned into pGEM-T Easy (Promega, USA), linearized with *Sal* I and *in vitro* transcribed using MAXIscript and m^{7}G(5′)ppp(5′)G cap analog (Ambion) in the presence and absence of [α-^{32}P]-UTP. Pre-mRNA was incubated with 30% of HeLa nuclear extracts (4C, Belgium), ATP (0.5 mM), creatine phosphate (20 mM), $MgCl_2$ (2 mM), 20 U of RNasin (Promega), PVA (2.6%) and 30% of buffer E (5% glycerol, 0.5 mM dithiothreitol, 0.2 mM EDTA, 100 mM KCl and 20 mM HEPES, pH 7.8) in a 25 μl volume at 30°C for 3 h. Reactions were stopped by adding proteinase K at a final concentration of 200 ng/μl. RNA was extracted with phenol, and precipitated with ethanol. Isotopically labeled RNA was loaded onto 5% PAGE (Fig. 3A). Non-labeled RNA was reverse-transcribed with M-MLV (H^-) RT (Promega, USA) and primer BP-RT. Two microlitres of the first-strand cDNA were used as a template for PCR with primers BP-F/BP-R at an annealing temperature of 56°C. Amplification primers were designed using the Oligo programme (version 6.63) according to an expected cDNA template. A single PCR product of 113 nucleotides was gel-purified and sequenced as described (92).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We wish to thank Gil Ast, Doug Black, Javier Caceres, Tom Cooper, Gideon Dreyfuss, Adrian Krainer, Steven Marsh, Akila Mayeda, Neema Mayor, Hiroto Okayama, Robin Reed, William Rigby, Zeev Ronai, Gavin Screaton, Chris Smith, Taki Tamura, Matt Wollerton and Alan Weiner for reagents and/or helpful discussions. This study was supported by the European Commission, the Value in People Award from the Wellcome Trust, the University of Southampton (I.V.), the Swiss National Science Foundation and the Canton of Geneva (A.K.).

REFERENCES

1. Black, D.L. (2000) Protein diversity from alternative splicing: a challenge for bioinformatics and postgenome biology. *Cell*, **103**, 367–370.
2. Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.
3. Sorek, R., Shamir, R. and Ast, G. (2004) How prevalent is functional alternative splicing in the human genome? *Trends Genet.*, **20**, 68–71.
4. Burge, C.B., Tuschl, T. and Sharp, P.A. (1999) Splicing of precursors to mRNAs by the spliceosome. In Gesteland, R.F., Cech, T.R. and Atkins J.F. (eds), *The RNA World.* Cold Spring Harbor Laboratory Press, New York, pp. 525–560.
5. Tacke, R. and Manley, J.L. (1999) Determinants of SR protein specificity. *Curr. Opin. Cell Biol.*, **11**, 358–362.
6. Graveley, B.R., Hertel, K.J. and Maniatis, T. (1999) SR proteins are 'locators' of the RNA splicing machinery. *Curr. Biol.*, **9**, R6–R7.
7. Cartegni, L., Chew, S.L. and Krainer, A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.*, **3**, 285–298.
8. Fairbrother, W.G., Yeh, R.F., Sharp, P.A. and Burge, C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science*, **297**, 1007–1013.
9. Zhang, X.H. and Chasin, L.A. (2004) Computational definition of sequence motifs governing constitutive exon splicing. *Genes Dev.*, **18**, 1241–1250.
10. Waldburger, J.M., Masternak, K., Muhlethaler-Mottet, A., Villard, J., Peretti, M., Landmann, S. and Reith, W. (2000) Lessons from the bare lymphocyte syndrome: molecular mechanisms regulating MHC class II expression. *Immunol. Rev.*, **178**, 148–165.
11. Briata, P., Radka, S.F., Sartoris, S. and Lee, J.S. (1989) Alternative splicing of HLA-DQB transcripts and secretion of HLA-DQ beta-chain proteins: allelic polymorphism in splicing and polyadenylylation sites. *Proc. Natl Acad. Sci. USA*, **86**, 1003–1007.
12. Aultman, D., Adamashvili, I., Yaturu, K., Langford, M., Gelder, F., Gautreaux, M., Ghali, G.E. and McDonald, J. (1999) Soluble HLA in human body fluids. *Hum. Immunol.*, **60**, 239–244.
13. Zavazava, N. and Kronke, M. (1996) Soluble HLA class I molecules induce apoptosis in alloreactive cytotoxic T lymphocytes. *Nat. Med.*, **2**, 1005–1010.
14. Rhode, P.R., Burkhardt, M., Jiao, J., Siddiqui, A.H., Huang, G.P. and Wong, H.C. (1996) Single-chain MHC class II molecules induce T cell activation and apoptosis. *J. Immunol.*, **157**, 4885–4891.
15. Weyand, C.M., Jendro, M. and Goronzy, J.J. (1991) Soluble HLA-DR molecules in patients with HLA class II versus class I associated disorders. *Autoimmunity*, **8**, 281–287.
16. Senju, S., Kimura, A., Yasunami, M., Kamikawaji, N., Yoshizumi, H., Nishimura, Y. and Sasazuki, T. (1992) Allele-specific expression of the cytoplasmic exon of HLA-DQB1 gene. *Immunogenetics*, **36**, 319–325.
17. Vogel, J., Hess, W.R. and Borner, T. (1997) Precise branch point mapping and quantification of splicing intermediates. *Nucleic Acids Res.*, **25**, 2030–2031.
18. Zhuang, Y.A., Goldstein, A.M. and Weiner, A.M. (1989) UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl Acad. Sci. USA*, **86**, 2752–2756.
19. Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C. and Markham, A.F. (1989) Analysis of any

point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.*, **17**, 2503–2516.
20. Berglund, J.A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. (1997) The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell.*, **89**, 781–787.
21. Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngninou-Molango, S., Sprangers, R., Zanier, K., Kramer, A. and Sattler, M. (2001) Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science*, **294**, 1098–1102.
22. Parker, R., Siliciano, P.G. and Guthrie, C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell*, **49**, 229–239.
23. Wu, J. and Manley, J.L. (1989) Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.*, **3**, 1553–1561.
24. Zhuang, Y. and Weiner, A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.*, **3**, 1545–1552.
25. Peled-Zehavi, H., Berglund, J.A., Rosbash, M. and Frankel, A.D. (2001) Recognition of RNA branch point sequences by the KH domain of splicing factor 1 (mammalian branch point binding protein) in a splicing factor complex. *Mol. Cell. Biol.*, **21**, 5232–5241.
26. Hartmuth, K. and Barta, A. (1988) Unusual branch point selection in processing of human growth hormone pre-mRNA. *Mol. Cell. Biol.*, **8**, 2011–2020.
27. Query, C.C., Strobel, S.A. and Sharp, P.A. (1996) Three recognition events at the branch-site adenine. *EMBO J.*, **15**, 1392–1402.
28. Li, M. and Pritchard, P.H. (2000) Characterization of the effects of mutations in the putative branchpoint sequence of intron 4 on the splicing within the human lecithin:cholesterol acyltransferase gene. *J. Biol. Chem.*, **275**, 18079–18084.
29. Abovich, N. and Rosbash, M. (1997) Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell*, **89**, 403–412.
30. Chavanas, S., Gache, Y., Vailly, J., Kanitakis, J., Pulkkinen, L., Uitto, J., Ortonne, J. and Meneguzzi, G. (1999) Splicing modulation of integrin beta4 pre-mRNA carrying a branch point mutation underlies epidermolysis bullosa with pyloric atresia undergoing spontaneous amelioration with ageing. *Hum. Mol. Genet.*, **8**, 2097–2105.
31. Kuivenhoven, J.A., Weibusch, H., Pritchard, P.H., Funke, H., Benne, R., Assmann, G. and Kastelein, J.J. (1996) An intronic mutation in a lariat branchpoint sequence is a direct cause of an inherited human disorder (fish-eye disease). *J. Clin. Invest.*, **98**, 358–364.
32. Burrows, N.P., Nicholls, A.C., Richards, A.J., Luccarini, C., Harrison, J.B., Yates, J.R. and Pope, F.M. (1998) A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehlers–Danlos syndrome type II in two British families. *Am. J. Hum. Genet.*, **63**, 390–398.
33. Hamilton, B.J., Genin, A., Cron, R.Q. and Rigby, W.F. (2003) Delineation of a novel pathway that regulates CD154 (CD40 ligand) expression. *Mol. Cell. Biol.*, **23**, 510–525.
34. Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q. and Krainer, A.R. (2003) ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res.*, **31**, 3568–3571.
35. Cavaloc, Y., Bourgeois, C.F., Kister, L. and Stevenin, J. (1999) The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. *RNA*, **5**, 468–483.
36. Schaal, T.D. and Maniatis, T. (1999) Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. *Mol. Cell. Biol.*, **19**, 1705–1719.
37. Gemignani, F., Sazani, P., Morcos, P. and Kole, R. (2002) Temperature-dependent splicing of beta-globin pre-mRNA. *Nucleic Acids Res.*, **30**, 4592–4598.
38. Touchman, J.W., D'Souza, I., Heckman, C.A., Zhou, R., Biggart, N.W. and Murphy, E.C., Jr (1995) Branchpoint and polypyrimidine tract mutations mediating the loss and partial recovery of the Moloney murine sarcoma virus MuSVts110 thermosensitive splicing phenotype. *J. Virol.*, **69**, 7724–7733.
39. Bruzik, J.P. and Steitz, J.A. (1990) Spliced leader RNA sequences can substitute for the essential 5′ end of U1 RNA during splicing in a mammalian *in vitro* system. *Cell*, **62**, 889–899.
40. Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866–1869.
41. Tarn, W.Y. and Steitz, J.A. (1994) SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. *Genes Dev.*, **8**, 2704–2717.
42. Hwang, D.Y. and Cohen, J.B. (1996) Base pairing at the 5′ splice site with U1 small nuclear RNA promotes splicing of the upstream intron but may be dispensable for slicing of the downstream intron. *Mol. Cell. Biol.*, **16**, 3012–3022.
43. Du, H. and Rosbash, M. (2001) Yeast U1 snRNP-pre-mRNA complex formation without U1snRNA-pre-mRNA base pairing. *RNA*, **7**, 133–142.
44. Lund, M. and Kjems, J. (2002) Defining a 5′ splice site by functional selection in the presence and absence of U1 snRNA 5′ end. *RNA*, **8**, 166–179.
45. Norton, P.A. (1994) Polypyrimidine tract sequences direct selection of alternative branch sites and influence protein binding. *Nucleic Acids Res.*, **22**, 3854–3860.
46. Berglund, J.A., Abovich, N. and Rosbash, M. (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.*, **12**, 858–867.
47. Shen, H., Kan, J.L. and Green, M.R. (2004) Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Mol. Cell*, **13**, 367–376.
48. MacMillan, A.M., Query, C.C., Allerson, C.R., Chen, S., Verdine, G.L. and Sharp, P.A. (1994) Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev.*, **8**, 3008–3020.
49. Gaur, R.K., Valcarcel, J. and Green, M.R. (1995) Sequential recognition of the pre-mRNA branch point by U2AF65 and a novel spliceosome-associated 28-kDa protein. *RNA*, **1**, 407–417.
50. Rose-John, S. and Heinrich, P.C. (1994) Soluble receptors for cytokines and growth factors: generation and biological function. *Biochem. J.*, **300**, 281–290.
51. Xing, Y., Xu, Q. and Lee, C. (2003) Widespread production of novel soluble protein isoforms by alternative splicing removal of transmembrane anchoring domains. *FEBS Lett.*, **555**, 572–578.
52. Cline, M.S., Shigeta, R., Wheeler, R.L., Siani-Rose, M.A., Kulp, D. and Loraine, A.E. (2004) The effects of alternative splicing on transmembrane proteins in the mouse genome. *Pac. Symp. Biocomput.*, **9**, 17–28.
53. Cosson, P. and Bonifacino, J.S. (1992) Role of transmembrane domain interactions in the assembly of class II MHC molecules. *Science*, **258**, 659–662.
54. Calafat, J., Nijenhuis, M., Janssen, H., Tulp, A., Dusseljee, S., Wubbolts, R. and Neefjes, J. (1994) Major histocompatibility complex class II molecules induce the formation of endocytic MIIC-like structures. *J. Cell Biol.*, **126**, 967–977.
55. Weber, D.A., Dao, C.T., Jun, J., Wigal, J.L. and Jensen, P.E. (2001) Transmembrane domain-mediated colocalization of HLA-DM and HLA-DR is required for optimal HLA-DM catalytic activity. *J. Immunol.*, **167**, 5167–5174.
56. Siebold, C., Hansen, B.E., Wyer, J.R., Harlos, K., Esnouf, R.E., Svejgaard, A., Bell, J.I., Strominger, J.L., Jones, E.Y. and Fugger, L. (2004) Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers strong susceptibility to narcolepsy. *Proc. Natl Acad. Sci. USA*, **101**, 1999–2004.
57. Strusberg, I., Mendelberg, R.C., Serra, H.A. and Strusberg, A.M. (2002) Influence of weather conditions on rheumatic pain. *J. Rheumatol.*, **29**, 335–338.
58. Challier, B., Urlacher, F., Vancon, G., Lemelle, I., Pourel, J. and Guillemin, F. (2001) Is quality of life affected by season and weather conditions in ankylosing spondylitis? *Clin. Exp. Rheumatol.*, **19**, 277–281.
59. Boniface, J.J., Reich, Z., Lyons, D.S. and Davis, M.M. (1999) Thermodynamics of T cell receptor binding to peptide–MHC: evidence for a general mechanism of molecular scanning. *Proc. Natl Acad. Sci. USA*, **96**, 11446–11451.
60. Ars, E., Serra, E., de la Luna, S., Estivill, X. and Lazaro, C. (2000) Cold shock induces the insertion of a cryptic exon in the neurofibromatosis type 1 (NF1) mRNA. *Nucleic Acids Res.*, **28**, 1307–1312.
61. Takechi, H., Hosokawa, N., Hirayoshi, K. and Nagata, K. (1994) Alternative 5′ splice site selection induced by heat shock. *Mol. Cell. Biol.*, **14**, 567–575.

62. Weil, D., D'Alessio, M., Ramirez, F., Steinmann, B., Wirtz, M.K., Glanville, R.W. and Hollister, D.W. (1989) Temperature-dependent expression of a collagen splicing defect in the fibroblasts of a patient with Ehlers–Danlos syndrome type VII. *J. Biol. Chem.*, **264**, 16804–16809.
63. Kosaki, A., Nelson, J. and Webster, N.J. (1998) Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. *J. Biol. Chem.*, **273**, 10331–10337.
64. Pagani, F., Stuani, C., Tzetis, M., Kanavakis, E., Efthymiadou, A., Doudounakis, S., Casals, T. and Baralle, F.E. (2003) New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum. Mol. Genet.*, **12**, 1111–1120.
65. Krainer, A.R., Conway, G.C. and Kozak, D. (1990) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.*, **4**, 1158–1171.
66. Mayeda, A. and Krainer, A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*, **68**, 365–375.
67. Kanopka, A., Muhlemann, O. and Akusjärvi, G. (1996) Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature*, **381**, 535–538.
68. Cook, C.R. and McNally, M.T. (1998) SR protein and snRNP requirements for assembly of the Rous sarcoma virus negative regulator of splicing complex *in vitro*. *Virology*, **242**, 211–220.
69. Wang, J. and Manley, J.L. (1995) Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing *in vivo* in diverse ways. *RNA*, **1**, 335–346.
70. Wang, J., Xiao, S.H. and Manley, J.L. (1998) Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. *Genes Dev.*, **12**, 2222–2233.
71. Wagner, E.J. and Garcia-Blanco, M.A. (2001) Polypyrimidine tract binding protein antagonizes exon definition. *Mol. Cell. Biol.*, **21**, 3281–3288.
72. Lou, H., Helfman, D.M., Gagel, R.F. and Berget, S.M. (1999) Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3′-terminal exon. *Mol. Cell. Biol.*, **19**, 78–85.
73. Flanagan, M., Liang, H. and Norton, P.A. (2003) Alternative splicing of fibronectin mRNAs in chondrosarcoma cells: role of far upstream intron sequences. *J. Cell Biochem.*, **90**, 709–718.
74. Jiang, Z.H., Zhang, W.J., Rao, Y. and Wu, J.Y. (1998) Regulation of Ich-1 pre-mRNA alternative splicing and apoptosis by mammalian splicing factors. *Proc. Natl Acad. Sci. USA*, **95**, 9155–9160.
75. Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. *Science*, **265**, 1706–1709.
76. Rooke, N., Markovtsov, V., Cagavi, E. and Black, D.L. (2003) Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. *Mol. Cell. Biol.*, **23**, 1874–1884.
77. Zhu, J., Mayeda, A. and Krainer, A.R. (2001) Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell*, **8**, 1351–1361.
78. Bai, Y., Lee, D., Yu, T. and Chasin, L.A. (1999) Control of 3′ splice site choice *in vivo* by ASF/SF2 and hnRNP A1. *Nucleic Acids Res.*, **27**, 1126–1134.
79. Fu, X.D., Mayeda, A., Maniatis, T. and Krainer, A.R. (1992) General splicing factors SF2 and SC35 have equivalent activities *in vitro*, and both affect alternative 5′ and 3′ splice site selection. *Proc. Natl Acad. Sci. USA*, **89**, 11224–11228.
80. Matunis, M.J., Xing, J. and Dreyfuss, G. (1994) The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization. *Nucleic Acids Res.*, **22**, 1059–1067.
81. Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L. and Krainer, A.R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.*, **138**, 225–238.
82. Nasim, M.T., Chernova, T.K., Chowdhury, H.M., Yue, B.G. and Eperon, I.C. (2003) HnRNP G and Tra2beta: opposite effects on splicing matched by antagonism in RNA binding. *Hum. Mol. Genet.*, **12**, 1337–1348.
83. Charlet, B.N., Logan, P., Singh, G. and Cooper, T.A. (2002) Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. *Mol. Cell*, **9**, 649–658.
84. Gromak, N., Matlin, A.J., Cooper, T.A. and Smith, C.W. (2003) Antagonistic regulation of alpha-actinin alternative splicing by CELF proteins and polypyrimidine tract binding protein. *RNA*, **9**, 443–456.
85. Labourier, E., Bourbon, H.M., Gallouzi, I.E., Fostier, M., Allemand, E. and Tazi, J. (1999) Antagonism between RSF1 and SR proteins for both splice-site recognition *in vitro* and *Drosophila* development. *Genes Dev.*, **13**, 740–753.
86. Singh, R., Valcárcel, J. and Green, M.R. (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*, **268**, 1173–1176.
87. Banerjee, H., Rahn, A., Davis, W. and Singh, R. (2003) Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. *RNA*, **9**, 88–99.
88. Roscigno, R.F., Weiner, M. and Garcia-Blanco, M.A. (1993) A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *J. Biol. Chem.*, **268**, 11222–11229.
89. Coolidge, C.J., Seely, R.J. and Patton, J.G. (1997) Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Res.*, **25**, 888–896.
90. Hunter, M., Bernard, R., Freitas, E., Boyer, A., Morar, B., Martins, I.J., Tournev, I., Jordanova, A., Guergelcheva, V., Ishpekova, B. *et al.* (2003) Mutation screening of the N-myc downstream-regulated gene 1 (NDRG1) in patients with Charcot–Marie-Tooth Disease. *Hum. Mutat.*, **22**, 129–135.
91. Vořechovský, I., Královičová, J., Laycock, M.D., Webster, A.D.B., Marsh, S.G.E., Madrigal, A. and Hammarström, L. (2001) Short tandem repeat (STR) haplotypes in *HLA*: an integrated 50-kb STR/linkage disequilibrium/gene map between the *RING3* and *HLA-B* genes and identification of STR haplotype diversification in the class III region. *Eur. J. Hum. Genet.*, **9**, 590–598.
92. Vořechovský, I., Luo, L., Dyer, M.J., Catovsky, D., Amlot, P.L., Yaxley, J.C., Foroni, L., Hammarström, L., Webster, A.D. and Yuille, M.A. (1997) Clustering of missense mutations in the ataxia-telangiectasia gene in a sporadic T-cell leukaemia. *Nat. Genet.*, **17**, 96–99.
93. Nei, M. (1987) *Molecular Evolutionary Genetics.* Columbia University Press, New York, NY.
94. Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X. and Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
95. Yamamoto, H., Tsukahara, K., Kanaoka, Y., Jinno, S. and Okayama, H. (1999) Isolation of a mammalian homologue of a fission yeast differentiation regulator. *Mol. Cell. Biol.*, **19**, 3829–3841.
96. Gil, A., Sharp, P.A., Jamison, S.F. and Garcia-Blanco, M.A. (1991) Characterization of cDNAs encoding the polypyrimidine tract-binding protein. *Genes Dev.*, **5**, 1224–1236.
97. Ghetti, A., Pinol-Roma, S., Michael, W.M., Morandi, C. and Dreyfuss, G. (1992) hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.*, **20**, 3671–3678.
98. Ladd, A.N., Charlet, N. and Cooper, T.A. (2001) The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol. Cell. Biol.*, **21**, 1285–1296.
99. Chua, K. and Reed, R. (1999) The RNA splicing factor hSlu7 is required for correct 3′ splice-site choice. *Nature*, **402**, 207–210.
100. Lei, H., Pospisilova, D., Lindblom, A. and Vořechovský, I. (2002) Dominant negative *ATM* mutations in breast cancer families. *J. Natl Cancer Inst.*, **94**, 951–952.
101. Mathews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
102. Robinson, J., Waller, M.J., Parham, P., de Groot, N., Bontrop, R., Kennedy, L.J., Stoehr, P. and Marsh, S.G. (2003) IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res.*, **31**, 311–314.
103. Mignot, E., Lin, L., Rogers, W., Honda, Y., Qiu, X., Lin, X., Okun, M., Hohjoh, H., Miki, T., Hsu, S. *et al.* (2001) Complex HLA–DR and –DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *Am. J. Hum. Genet.*, **68**, 686–699.
104. Rosenthal, A., Jouet, M. and Kenwrick, S. (1992) Aberrant splicing of neural cell adhesion molecule L1 mRNA in a family with X-linked hydrocephalus. *Nat. Genet.*, **2**, 107–112.

105. Brand, K., Dugi, K.A., Brunzell, J.D., Nevin, D.N. and Santamarina-Fojo, S. (1996) A novel A→G mutation in intron I of the hepatic lipase gene leads to alternative splicing resulting in enzyme deficiency. *J. Lipid Res.*, **37**, 1213–1223.
106. Putnam, E.A., Park, E.S., Aalfs, C.M., Hennekam, R.C. and Milewicz, D.M. (1997) Parental somatic and germ-line mosaicism for a FBN2 mutation and analysis of FBN2 transcript levels in dermal fibroblasts. *Am. J. Hum. Genet.*, **60**, 818–827.
107. Fujimaru, M., Tanaka, A., Choeh, K., Wakamatsu, N., Sakuraba, H. and Isshiki, G. (1998) Two mutations remote from an exon/intron junction in the beta-hexosaminidase beta-subunit gene affect 3′-splice site selection and cause Sandhoff disease. *Hum. Genet.*, **103**, 462–469.
108. Khan, S.G., Metin, A., Gozukara, E., Inui, H., Shahlavi, T., Muniz-Medina, V., Baker, C.C., Ueda, T., Aiken, J.R., Schneider, T.D. *et al.* (2003) Two essential splice lariat branchpoint sequences in one intron in a xeroderma pigmentosum DNA repair gene: mutations result in reduced XPC mRNA levels that correlate with cancer risk. *Hum. Mol. Genet.*, **13**, 343–352.
109. Ketterling, R.P., Drost, J.B., Scaringe, W.A., Liao, D.Z., Liu, J.Z., Kasper, C.K. and Sommer, S.S. (1999) Reported *in vivo* splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations. *Hum. Mutat.*, **13**, 221–231.
110. Mayer, K., Ballhausen, W., Leistner, W. and Rott, H. (2000) Three novel types of splicing aberrations in the tuberous sclerosis *TSC2* gene caused by mutations apart from splice consensus sequences. *Biochim. Biophys. Acta*, **1502**, 495–507.
111. Janssen, R.J., Wevers, R.A., Haussler, M., Luyten, J.A., Steenbergen-Spanjers, G.C., Hoffmann, G.F., Nagatsu, T. and Van den Heuvel, L.P. (2000) A branch site mutation leading to aberrant splicing of the human tyrosine hydroxylase gene in a child with a severe extrapyramidal movement disorder. *Ann. Hum. Genet.*, **64**, 375–382.